

Tackling Data Science Challenges in the Exascale Computing Project



Lori Diachin, Lawrence Livermore National Laboratory (LLNL)
Deputy Director, Exascale Computing Project
www.exascaleproject.org

NYSDES 2021

October 26, 2021



The Exascale Computing Project (ECP) enables US revolutions in technology development; scientific discovery; healthcare; energy, economic, and national security

ECP
mission

Develop exascale-ready applications and solutions that address currently intractable problems of strategic importance and national interest.

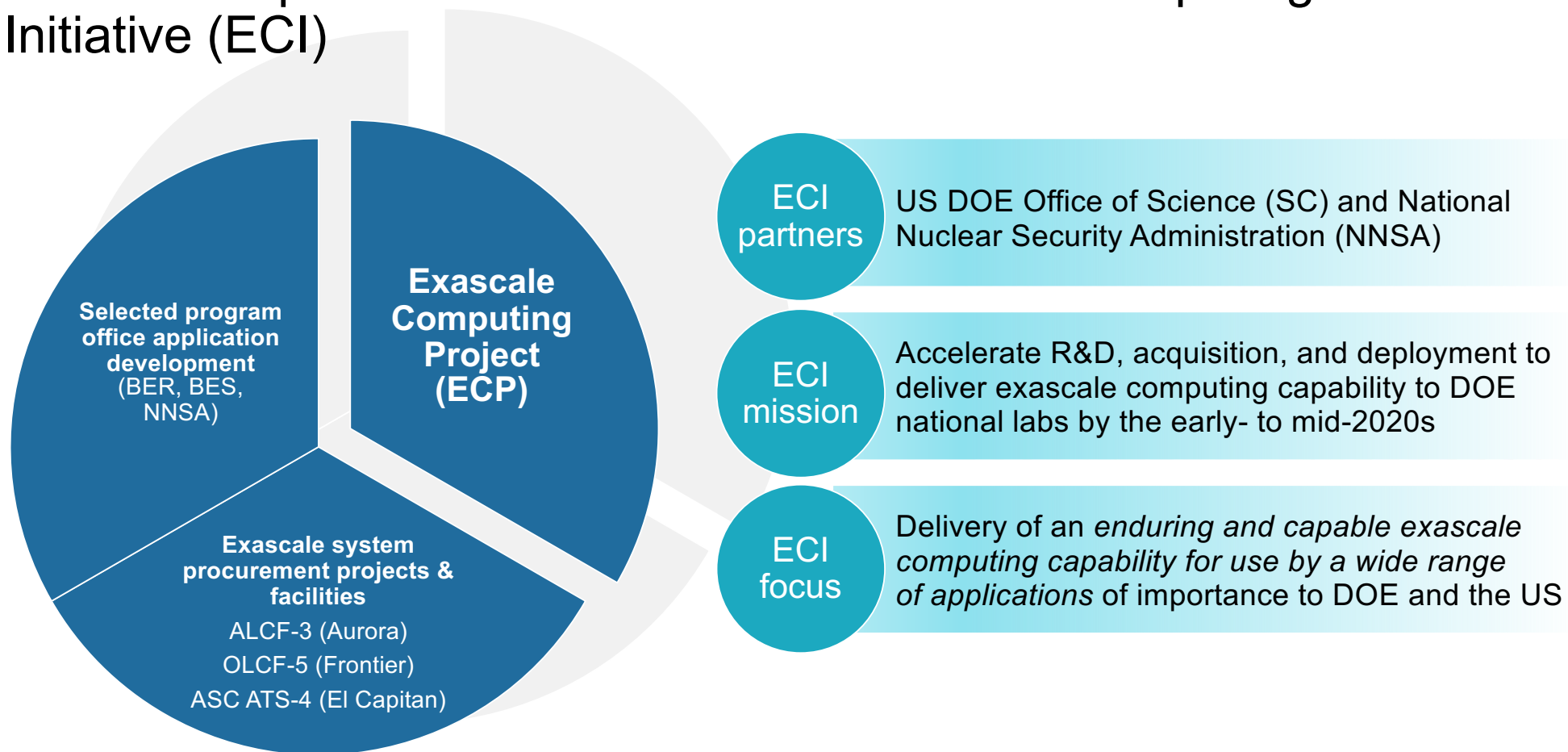
Create and deploy an expanded and vertically integrated software stack on DOE HPC exascale and pre-exascale systems, defining the enduring US exascale ecosystem.

Deliver **US HPC vendor technology advances** and **deploy ECP products** to DOE HPC pre-exascale and exascale systems.

ECP
vision

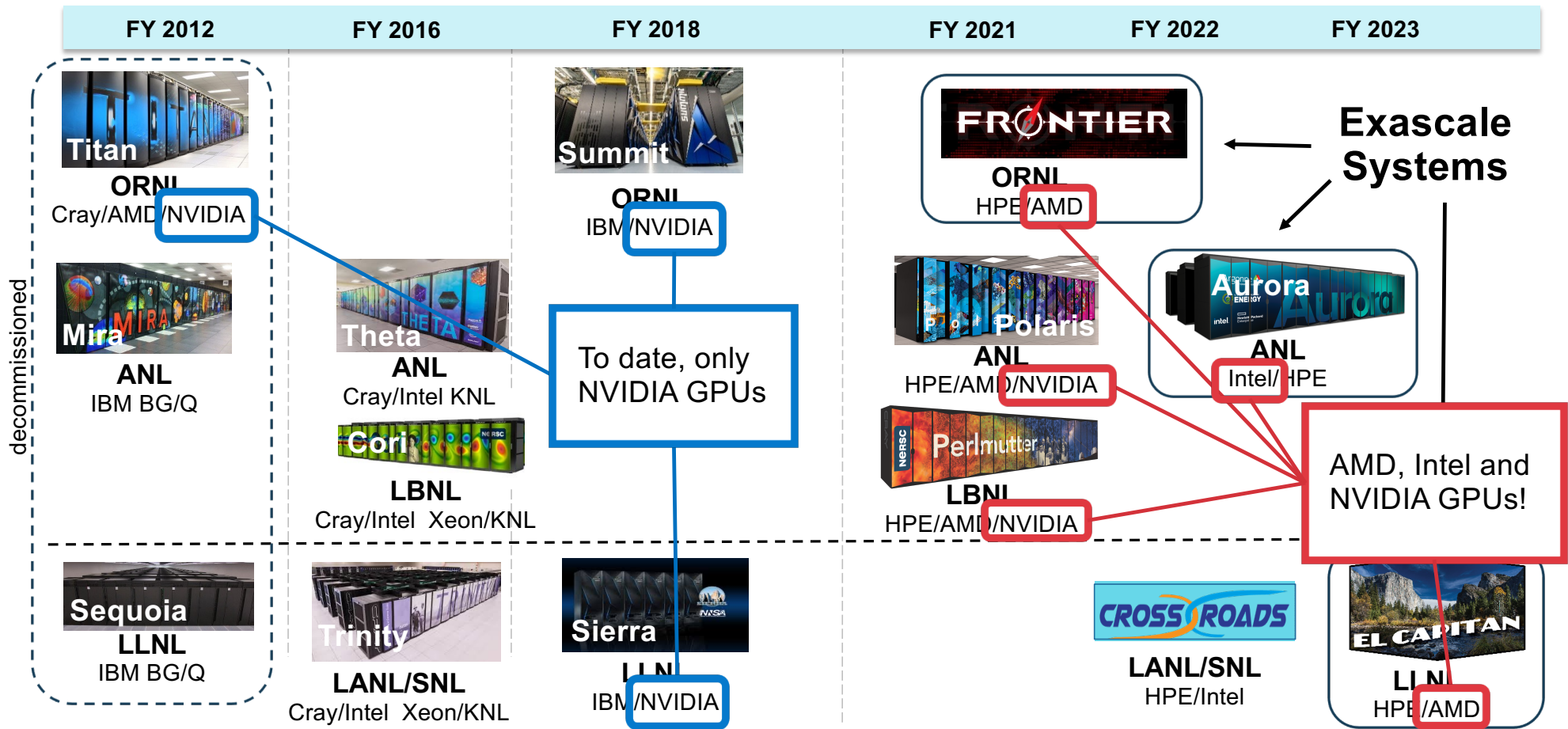
Deliver **exascale simulation and data science innovations and solutions to national problems** that enhance US economic competitiveness, change our quality of life, and strengthen our national security.

The ECP is part of the broader DOE Exascale Computing Initiative (ECI)



Three Major Components of the ECI

DOE HPC Roadmap to Exascale Systems



These systems will present significant challenges and opportunities for several aspects of data science and analytics

- **Challenges**

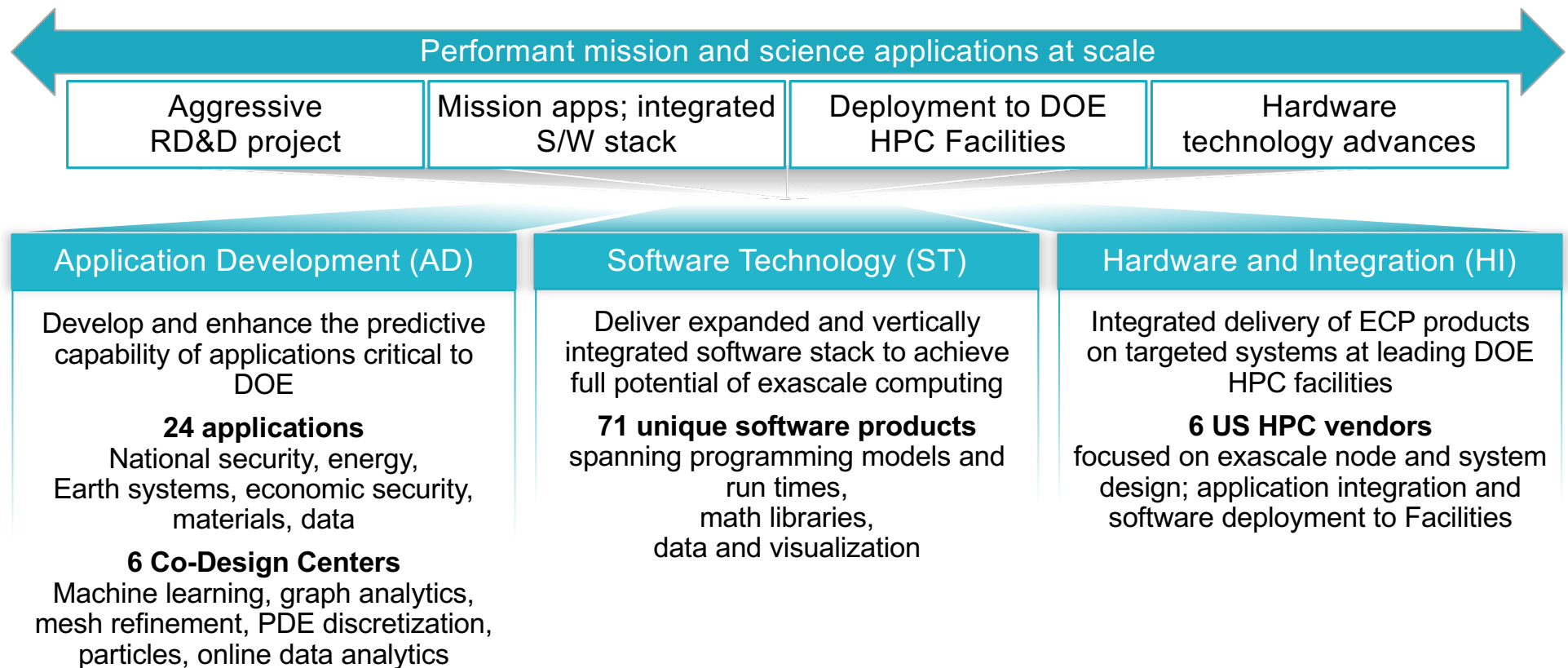
- Increasing discrepancy between compute and IO speeds
- Extremely large data sets generated by scientific simulations
- GPU architectures are not always well aligned with the tasks associated with data analysis
- Increasing complexity of scientific workflows including multiple simulation codes, AI/ML, analysis

- **Opportunities**

- Large-scale data analysis through tighter coupling of experimental and HPC facilities
- Use of AI/ML to improve simulations
- Use of HPC and AI/ML for control of experimental facilities
- Community-driven and adopted solutions

The ECP is addressing these challenges and opportunities through a variety of projects

The ECP is organized into three different technical focus areas



Addressing data science challenges happens primarily in the ST and Co-Design areas and progress is highlighted in applications

Software Technologies

- ALPINE – data analytics and visualization tools (PI: J Ahrens, LANL)
- ExaWorks – software development toolkits for complex workflows (PI: D. Laney, LLNL)

Co-Design Centers

- CoDAR – compression and data analytics (PI: I Foster, ANL)
- ExaLearn – machine learning (PI: F Alexander, BNL)

Applications

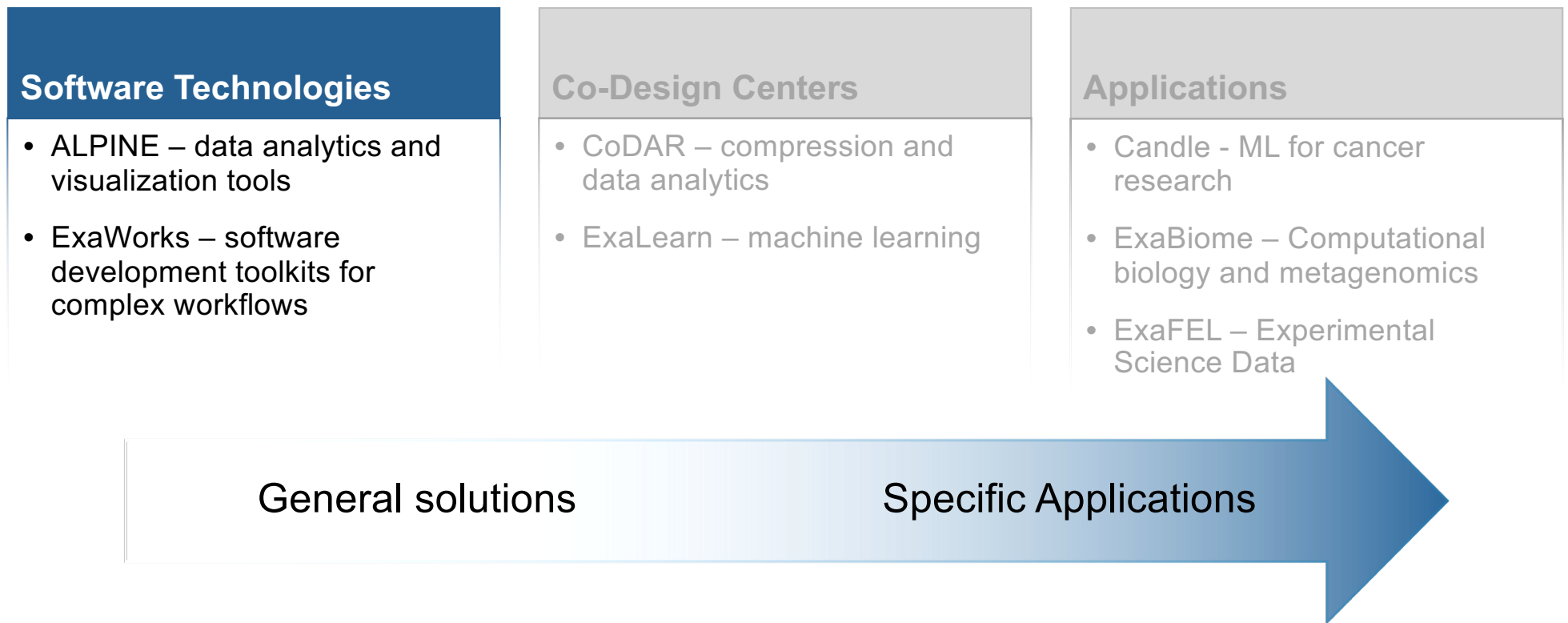
- Candle - ML for cancer research (PI: R Stevens, ANL)
- ExaBiome – Computational biology and metagenomics (PI: K Yelick, LBNL)
- ExaFEL – Experimental Science Data (PI: A. Perazzo, SLAC)

General solutions

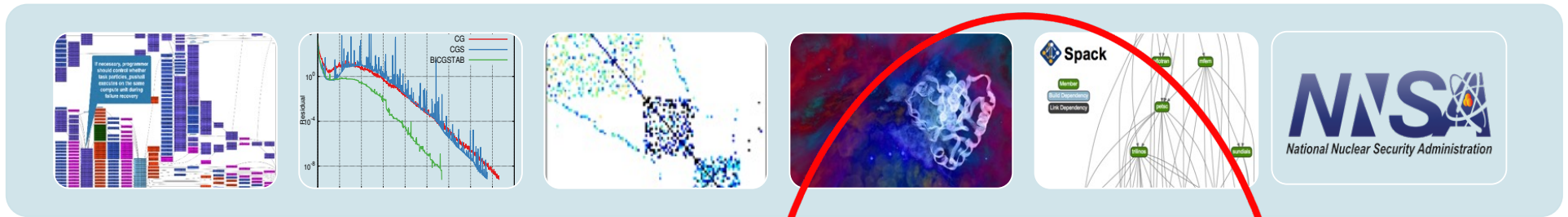
Specific Applications

Many many thanks to the great research teams that are conducting this work!

There are several key applications and technologies that highlight data science challenges and solutions at exascale



ECP's Software Technology effort spans six technical areas



Programming Models & Runtimes

- Enhance and get ready for exascale the widely used MPI and OpenMP programming models (hybrid programming models, deep memory copies)
- Development of performance portability tools (e.g. Kokkos and Raja)
- Support alternate models for potential benefits and risk mitigation: PGAS (UPC++/GASNet), task-based models (Legion, PARSEC)
- Libraries for deep memory hierarchy and power management

Development Tools

- Continued, multifaceted capabilities in portable, open-source LLVM compiler ecosystem to support expected ECP architectures, including support for F18
- Performance analysis tools that accommodate new architectures, programming models, e.g., PAPI, Tau

Math Libraries

- Linear algebra, iterative linear solvers, direct linear solvers, integrators and nonlinear solvers, optimization, FFTs, etc
- Performance on new node architectures; extreme strong scalability
- Advanced algorithms for multi-physics, multiscale simulation and outer-loop analysis
- Increasing quality, interoperability, complementarity of math libraries

Data and Visualization

- I/O via the HDF5 API
- **Insightful, memory-efficient in-situ visualization and analysis – Data reduction via scientific data compression**
- Checkpoint restart

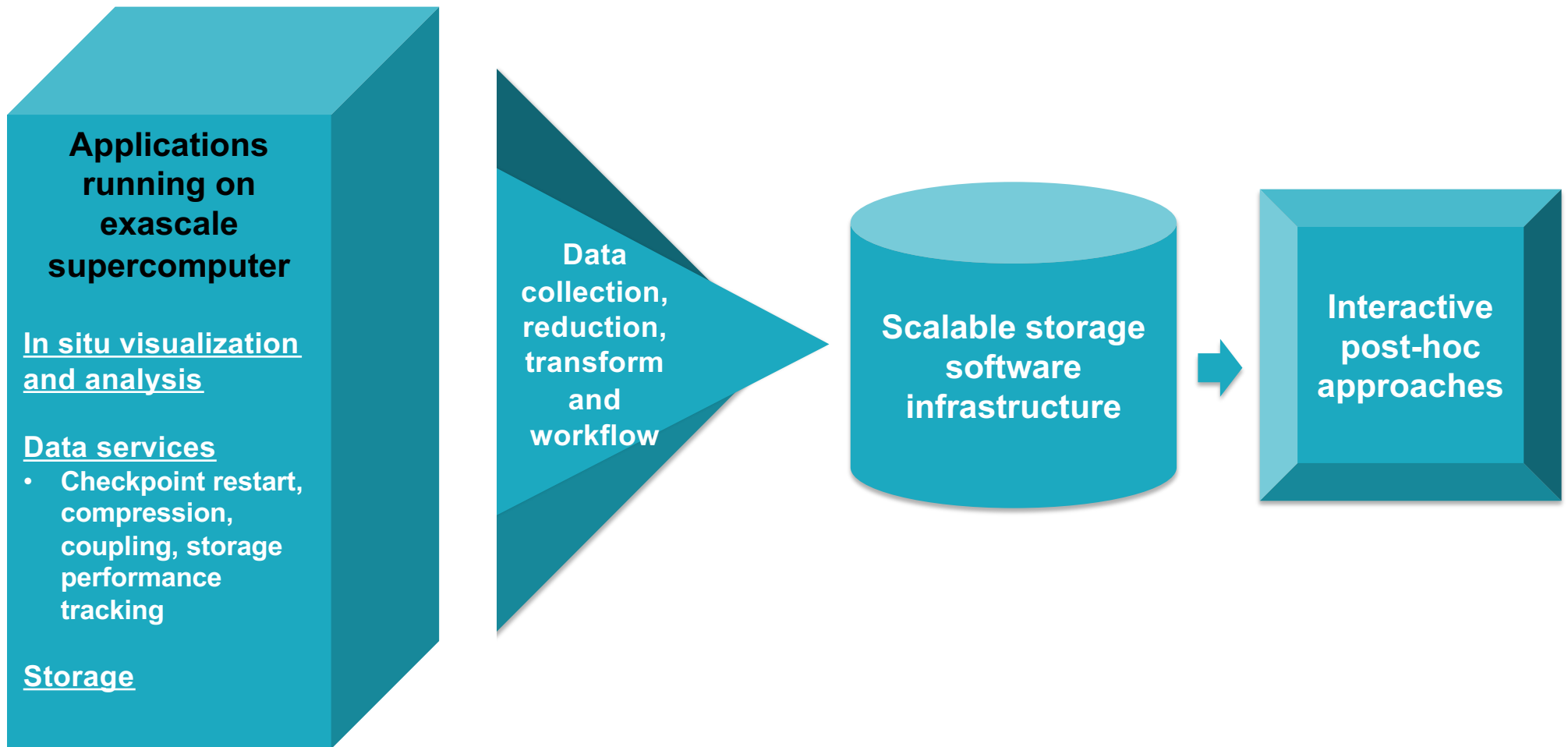
Software Ecosystem

- Develop features in Spack necessary to support all ST products in E4S, and the AD projects that adopt it
- Development of Spack stacks for reproducible turnkey deployment of large collections of software
- **Workflow software development toolkits**
- Regular E4S releases of the ST software stack and SDKs with regular integration of new ST products

NNSA ST

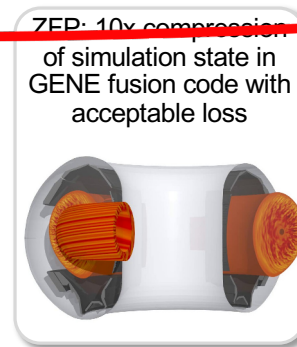
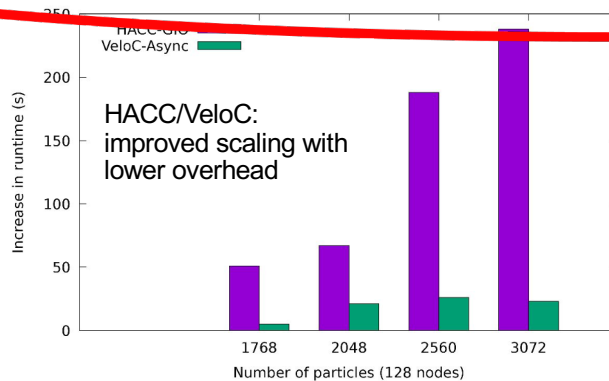
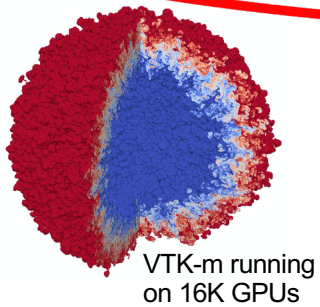
- Open source NNSA Software projects
- Projects that have both mission role and open science role
- Major technical areas: New programming abstractions, math libraries, data and viz libraries
- Cover most ST technology areas
- Subject to the same planning, reporting and review processes

Overview of ECP Data and Visualization Portfolio

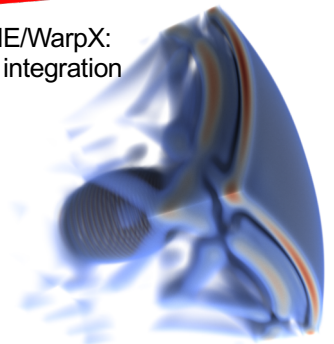


ECP Data and Visualization

Project Name	PI Name, Inst	Short Description/Objective
Data & Vis SDK	Chuck Atkins, Kitware	Support the deployment, testing and usage of ECP Data and Visualization packages.
ADIOS	Scott Klasky, ORNL	Supports efficient I/O and code coupling services.
DataLib	Rob Ross, ANL	Supports efficient I/O including MPI-IO (ROMIO) and PNetCDF, as well as I/O monitoring (Darshan) and data services (Mochi)
ECP/VTK-m	Ken Moreland, SNL	Comprehensive effort to provide VTK-based scientific visualization software that supports shared memory parallelism.
VeloC	Franck Cappello, ANL	VeloC-SZ develops two software products: VeloC (checkpoint restart) and SZ (lossy compression with strict error bounds).
ExaIO	Suren Byrta, LBNL	Efficient system topology and storage hierarchy-aware HDF5 and Unify (node local) parallel I/O libraries
ALPINE	James Ahrens, LANL	Deliver in situ visualization and analysis algorithms, infrastructure (ALPINE) and data reduction of floating-point arrays for reduced memory, communication, I/O, and offline storage via advanced data compression (ZFP).



ALPINE/WarpX: in situ integration



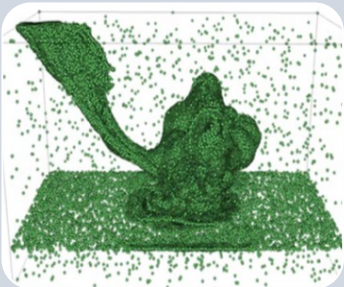
ALPINE Overview

Problem: Compute power growing faster than I/O bandwidth; requires running visualization *in situ* (during the simulation) to reduce I/O in scientific computing applications

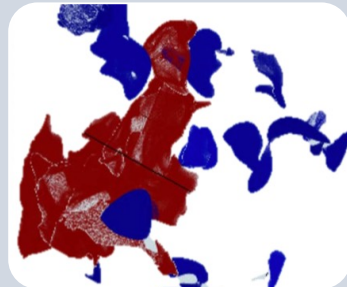
ALPINE delivers in situ visualization and analysis **algorithms** and **infrastructure**.

- Automated in situ massive data reduction **algorithms**: statistical feature exploration, adaptive sampling, topology, task-based feature detection, optimal viewpoint, Lagrangian flow, ...
 - **Pros**: No or greatly reduced I/O vs post-hoc processing, access to all data, computational power is readily available
 - **Cons**: Must know what you are looking for a priori, increasing complexity, memory and network constraints
- A portable, scalable, performant in situ **infrastructure** called Ascent

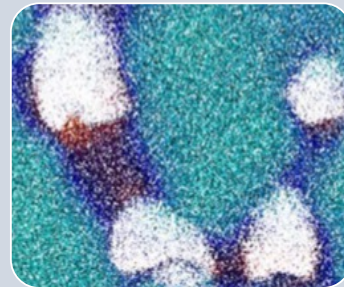
ALPINE algorithms help ECP scientists make the most of extreme datasets



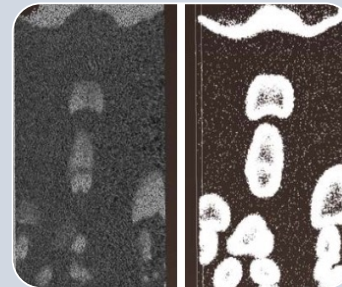
Data Driven Sampling enables probabilistic identification of interesting regions in the data automatically, prioritizing important regions.



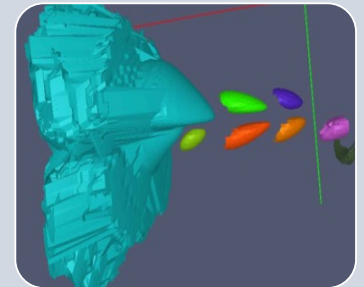
Task-based hierarchical feature extraction algorithm based on segmented merger tree. The algorithm is implemented using a multi-runtime abstraction layer, BabelFlow, which can be used to execute arbitrary analysis and visualization dataflows using different task-based runtime systems.



In situ **statistical feature detection and characterization** detects user specified features in particle data sets using statistical data modeling and probabilistic similarity measures



Rotation invariant pattern detection is used to identify bubble edges, reducing data by saving only areas of interest.

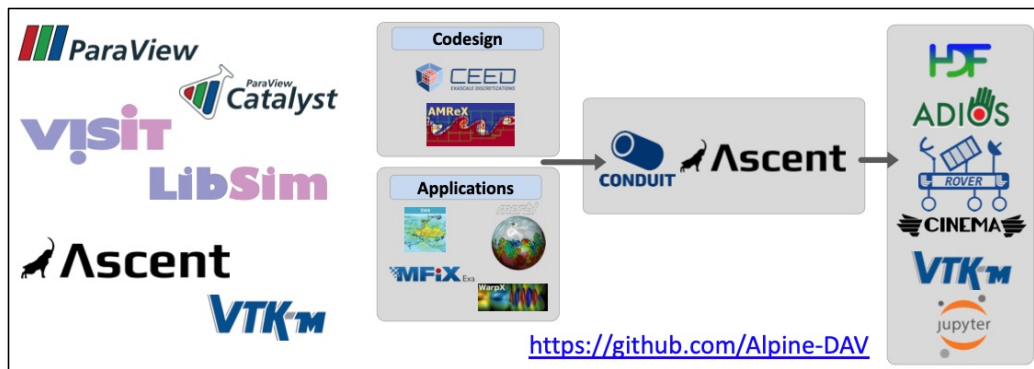


Topological analysis is used to identify most relevant contours and create isosurface visualizations *in situ*; saving only resulting images for post hoc analysis.

ALPINE Infrastructure builds on well known and widely used visualization infrastructure

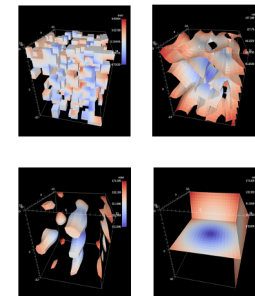
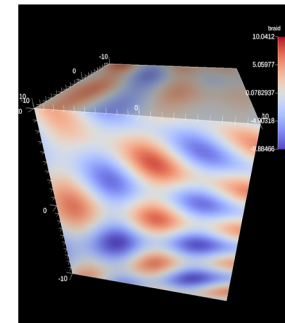
ParaView & VisIt and their in situ libraries (Catalyst and LibSim)

- Long-term DOE investments; two commonly used software packages for large-scale visualization and analysis within DOE SC and DOE NNSA
- ALPINE developers are ParaView and VisIt developers
- ALPINE will deliver in situ functionality leveraging ECP/VTK-m for performance and portability



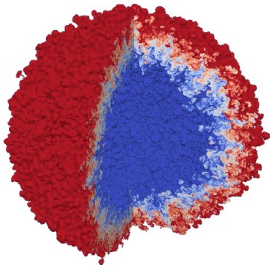
Ascent

- An API that links into a simulation code:
 - Making Pictures -- "scenes"
 - Transforming Data -- "filter" or "pipeline"
 - Capturing Data -- "publish"
- Efficient distributed memory (MPI) and many-core (CUDA or OpenMP) execution; lower memory requirements than many tools
- Provides infrastructure for applications to access Vis, Storage and Service software technologies



Supports common visualization use cases including pseudocolor, threshold, contours, isovolumes, slices, etc

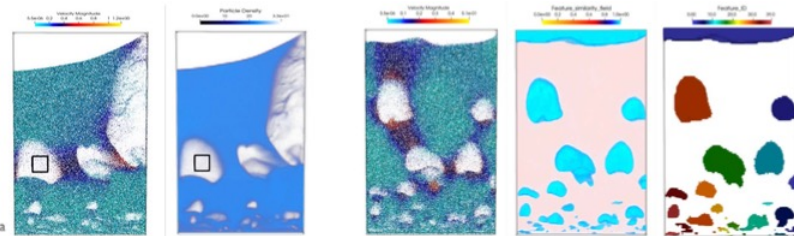
ALPINE has enabled new science across a number of domains



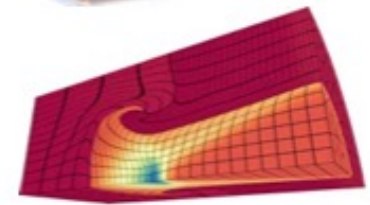
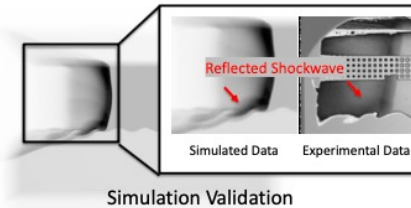
*Idealized Inertial Confinement
Fusion Simulation*

Rendering at scale using Ascent

- The 97.8 billion element ICF Rayleigh Taylor mixing simulation ran across 16,384 GPUs on 4,096 Nodes
- Time-varying evolution of the mixing was visualized in-situ using Ascent, also leveraged 16,384 GPUs
- Ascent used VTK-m to run visualization algorithms on the GPUs



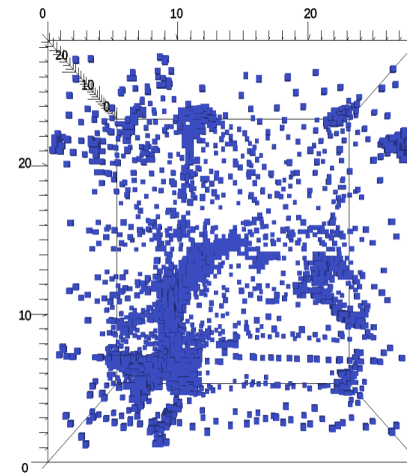
Identifying “voids” or bubble in MFX-Exa particle data **using statistical feature detection**. L to R: (1) Void selection: box shows region of interest with low particle density. (2) Voids as seen in density field. (3) Voids in particle field. (4) Feature similarity field – how similar to region of interest. (5) Voids are isolated by thresholding for analysis and characterization.



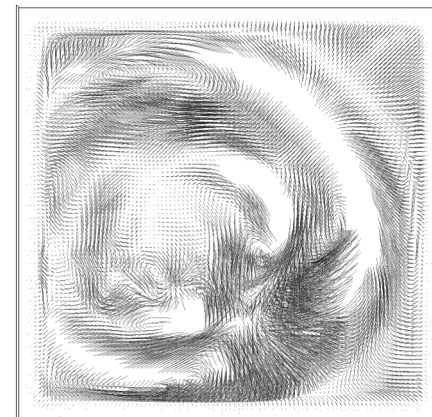
Examples of Rover and Ascent integration to support **in situ volume rendering** and simulated radiography for the MARBL simulation code. Used for comparison of simulation and experimental data. Native support for higher order meshes is in beta development

These capabilities are built on vtk-m; a toolkit of scientific visualization algorithms for emerging processor architectures

- Supports the fine-grained concurrency for data analysis and visualization algorithms required to drive extreme scale computing
- Provides abstract models for data and execution that can be applied to a variety of algorithms across many different processor architectures.
- Redevelop, implement, and support necessary visualization algorithms on many-core architectures
- Leverages performance portability tools such as Kokkos



Sampling of an ExaSky cosmology simulation run on Summit.

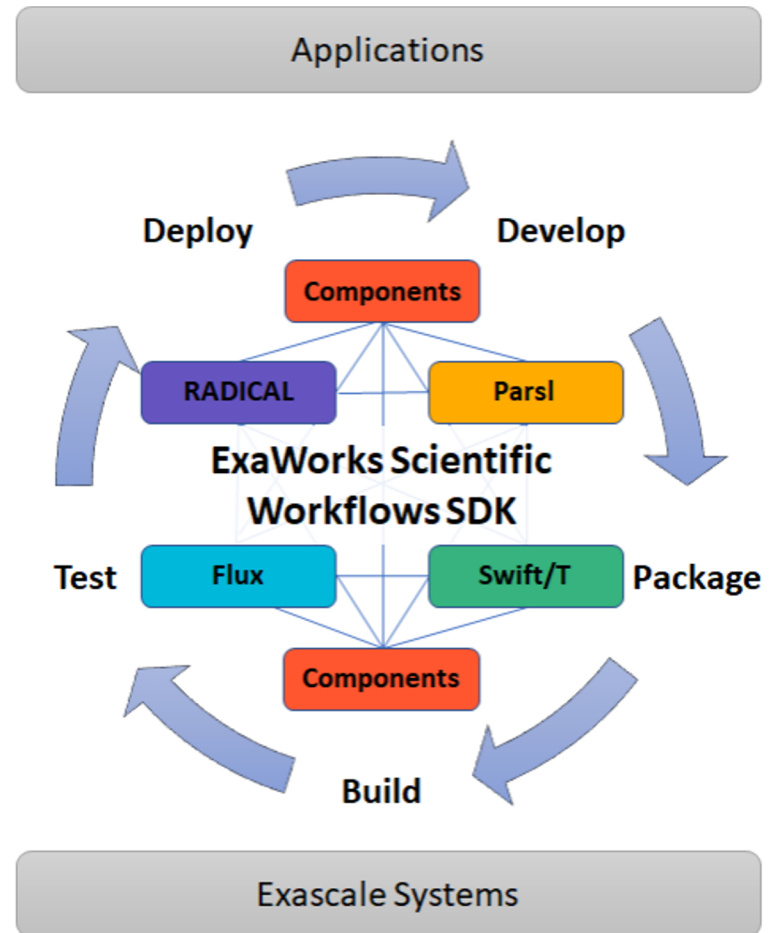


Lagrangian flow extracted by VTK-m from an EQSIM seismic simulation run on Summit.

The ExaWorks SDK provides a set of robust, scalable, and performance-driven components and APIs to users

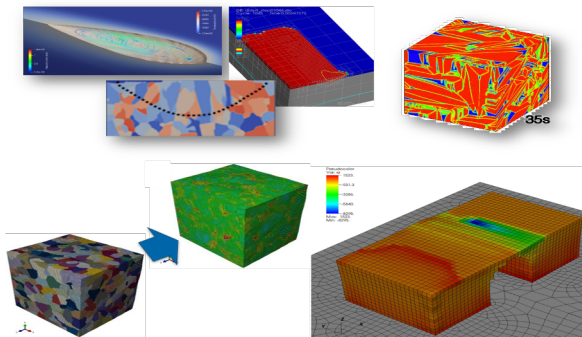
- Implemented via community-based approach
- Initial SDK technologies are used in ECP and have proven scalability
- Community development activities
 - Definition of inclusion policies and common packaging requirements (based on E4S)
 - Continuous integration & deployment infrastructure to foster adoption and community
 - Working toward integration across technologies
 - Development of common API's through an open community-based design process
- Initial alpha-release in October:

 <https://exaworks.org/>; <https://github.com/ExaWorks>
EXASCALE COMPUTING PROJECT

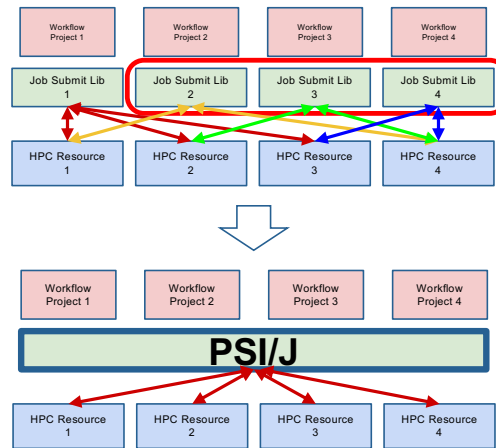


The ExaWorks project is engaging the workflow community and DOE facilities and impacting applications

Application Impact: 3 of 4 Gordon bell Covid-19 finalists and three ECP applications leveraging ExaWorks Technologies



PSI/J: a common portable job submission API



Specification: <https://exaworks.org/job-api-spec>

Reference Python binding:

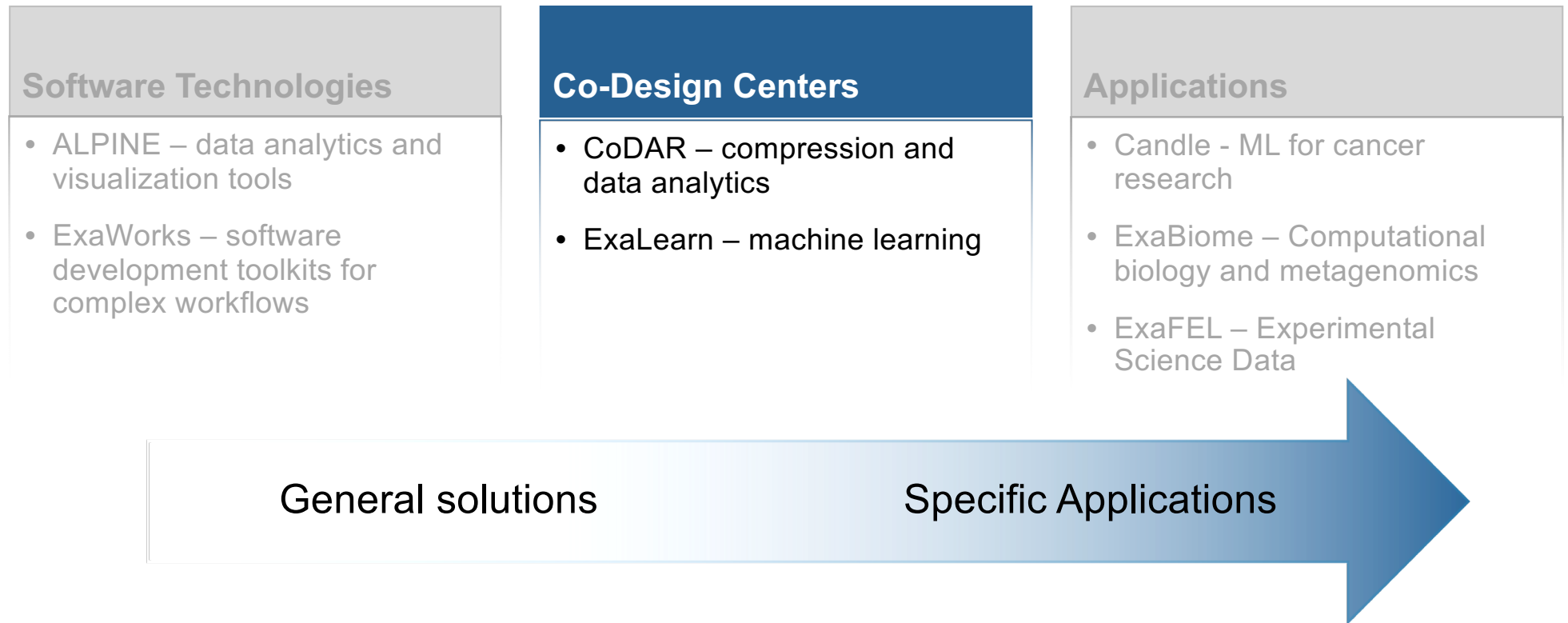
<https://github.com/ExaWorks/psi-j-python>

Workflows Community Summits:
January 2021: 45 workflow leaders representing 27+ workflow systems
April 2021: 75+ developers and users

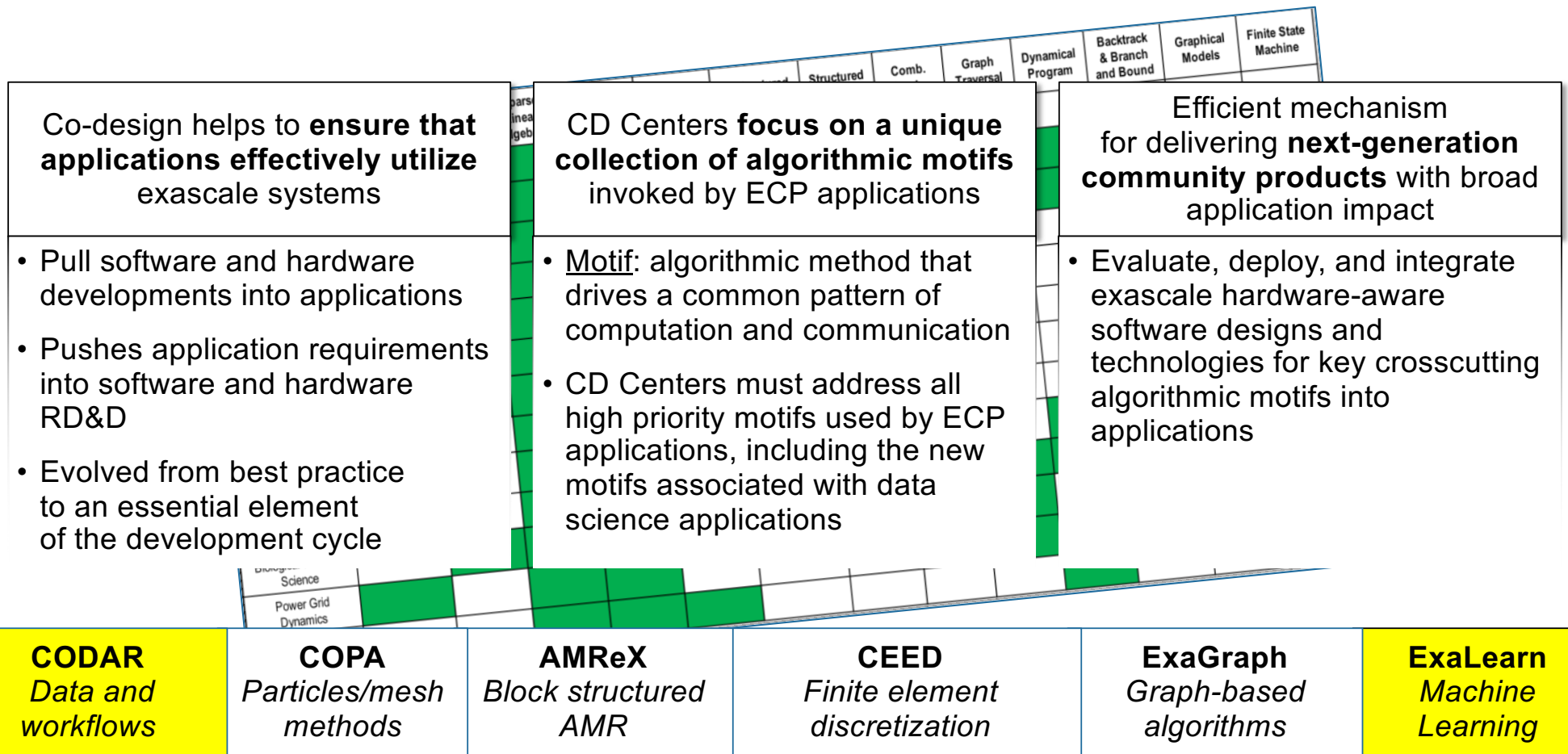


November 2021: ExaWorks & NSF WorkflowsRI Facilities Workshop

The ECP Co-Design portfolio sits one step closer to ECP applications

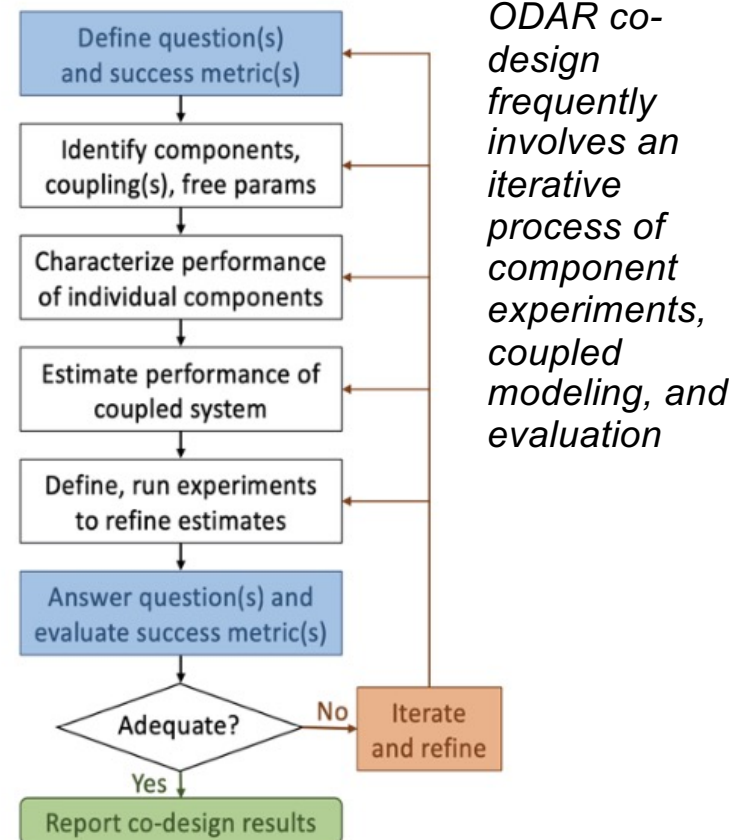


Co-design centers address computational motifs common to multiple application projects

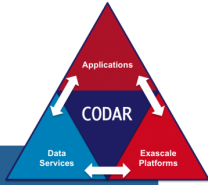


CODAR addresses the challenge of a 10^{12} disparity between computing speed and I/O speed

- Center for Online Data Analysis and Reduction (CODAR) Mission
 - Identify the best data analysis and reduction algorithms for different application classes, in terms of speed, accuracy, and resource requirements.
 - Quantify tradeoffs in data analysis accuracy, resource needs, and overall application performance among various data reduction methods. How do these tradeoffs vary with exascale hardware and software choices?
 - Effectively orchestrate online data analysis and reduction to maximize performance and provide flexibility
- A key challenge: GPUs are not that helpful for internode communication bottlenecks
- Key ECP customers: WDMapp, CANDLE, NWChemEx



Key technologies for CODAR success



SZ Lossy Compression Framework



A software framework for data and compression error analysis

Features:

- First modular, highly customizable/tunable lossy compressor (each compression stage can use different algorithms)
- Multiple compression error controls: Point wise error bounds (absolute, relative), statistical error bounds (PSNR). Dedicated error control for some applications.
- Integrated into ADIOS, HDF5, NetCDF, I/O libraries

<https://szcompressor.org>



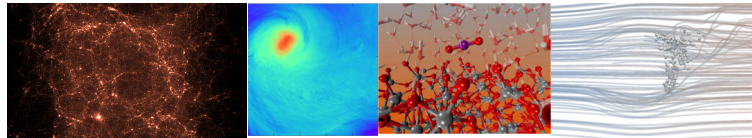
MGARD Lossy Compressor

MGARD: Multigrid Adaptive Reduction of Data, based on orthogonal projection

Features:

- Works on tensor product grids and cell-based data (triangles and tetrahedra)
- Can compress while preserving analysis results
- Adaptive in both space and time

<https://github.com/CODARcode/MGARD>



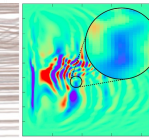
Z-checker

A compression framework for scientific data

Features:

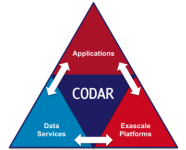
- Analyze initial data properties: Entropy, statistics
- Analyze the compression error with >30 metrics by comparing the initial data with decompressed data
- Produce a pdf report and visualization
- Read ADIOS, HDF5, NetCDF, plain binary stream of data

<https://github.com/CODARcode/Z-checker>



Tools used in Cosmology, Quantum Chemistry, Molecular dynamics, etc.

One of the unique opportunities ECP has is to provide benchmarks for scientific data reduction



A community repository providing reference scientific datasets (mostly ECP application), compressors (lossless and lossy), and error analysis tools

Features:

- Collection of representative datasets from ECP and other applications via direct communication with application developers and users
- Storage of the datasets on the Petrel server at Argonne with Terabytes of storage capacity
- Fast access to the datasets using Globus and GridFTP
- Access open to public

Co-design aspects:

- A critical tool to co-design lossy compression schemes for ECP applications
- Existing compressors can be optimized/tunes based on these datasets
- New compression schemes are developed with insurance that the dataset are coming from actual users and application developers

Impact:

- Already recognized as a reference source of information for the developers of lossy and lossless compressors for scientific data



<https://sdrbench.github.io/>

Scientific Data Reduction Benchmarks

This site has been established as part of the [ECP CODAR](#) project.

This site provides reference scientific datasets, data reduction techniques, error metrics, error controls and error assessment tools for users and developers of scientific data reduction techniques.

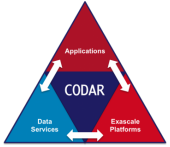
Important: when publishing results from one or more datasets presented in this webpage, make sure to:

- If the purpose is comparisons between compressors listed in this page, make sure to contact the compressor authors to get the correct compressor configuration according to each dataset and each comparison metrics
- Reference: Scientific Data Reduction Benchmarks (authors are the contributors/maintainers), <https://sdrbench.github.io/> and
- Acknowledge the source of the dataset you used and the DOE NNSA ECP project and the ECP CODAR project.
- Check the condition of publications (some dataset sources request prior check)

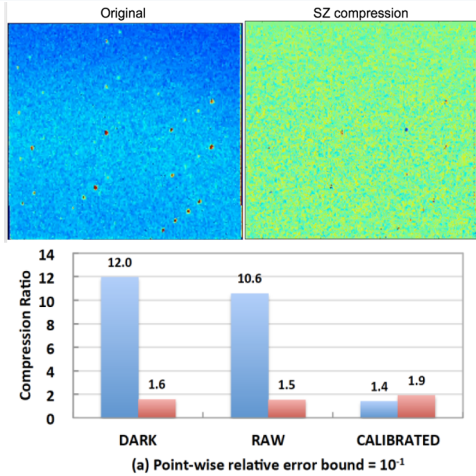
Data sets:

Name	Type	Format	Size (data)	Link
CESM-ATM <small>Source: Mark Taylor (SNL)</small>	Climate simulation	79 fields, 2D, 1800 x 3600, single precision, binary	1.47 GB	Dataset Metadata
EXAALT <small>Source: EXAALT team This dataset has been approved for unlimited release by Los Alamos National Laboratory and has been assigned LA-UR-18-25670.</small>	Molecular dynamics simulation	6 fields: x,y,z,vx,vy,vz, Each field stored separately, Single precision, Binary, Little-endian	60 MB	Dataset Metadata
EXAFEL <small>Source: LCLS</small>	Images from the LCLS instrument	2D, Single precision HDF5 and binary	51 MB	Dataset Metadata
HACC <small>Source: HACC team (ECP EXASKY)</small>	Cosmology: particle simulation	1 snapshot: 6 fields x,y,z,vx,vy,vz Each field stored separately, Single precision, Binary, Little-endian	19 GB 5 GB	Dataset Metadata Dataset Metadata
NYX <small>Source: Lukic et al. 'methods: numerical, intergalactic medium, quasars: absorption lines, large-scale structure of universe', Journal of Monthly Notices of Royal Astronomical Society.</small>	Cosmology: Adaptive mesh hydrodynamics + N-body cosmological simulation	6 fields, 3D, 512 x 512 x 512 Single precision, Binary, Little-endian	2.7 GB	Dataset Metadata

CODAR is targeting several online data analysis and reduction motifs for use in applications

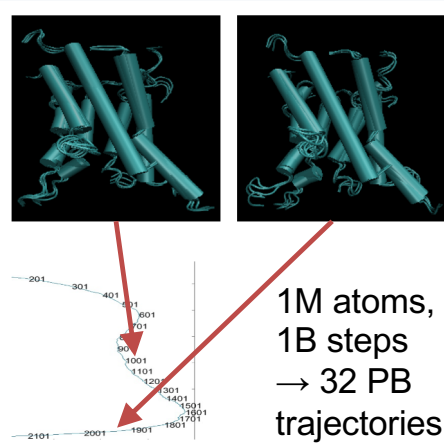


Online reduction



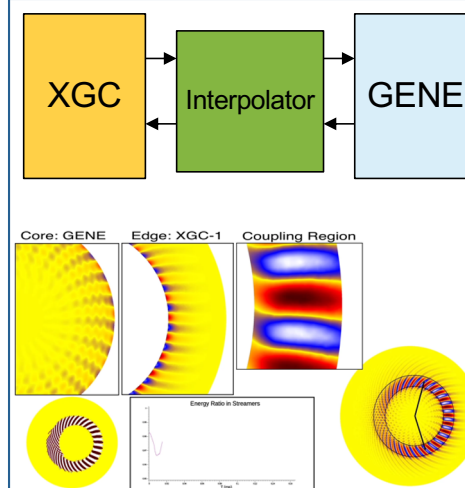
ExaFEL:
X-ray laser imaging

Online analysis



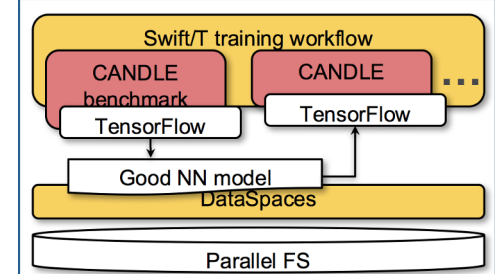
NWChemEx:
Molecular dynamics

Online coupling



WDMApp: Fusion
whole device model

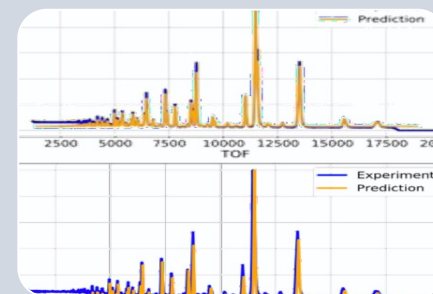
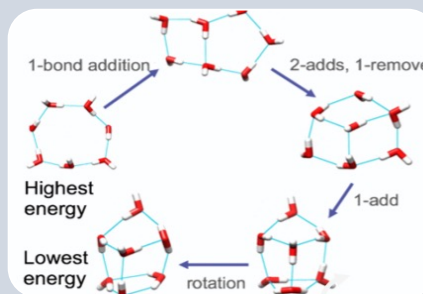
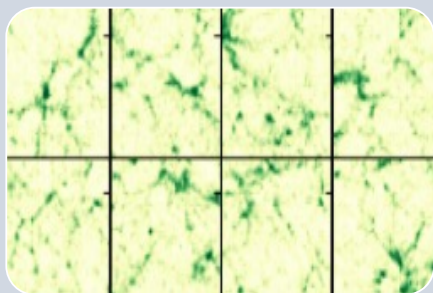
Online aggregation



Hyperparam. optimization:
 10^3 – 10^6 training runs, each
fitting many parameters

CANDLE:
Cancer deep learning

The ExaLearn co-design center is focused on developing machine learning technologies along four different application pillars



Surrogates replace computationally expensive simulations through machine learning (ML), cheaply

Method Used: Generative adversarial networks (GAN) and hybrid autoencoders.

Example Applications: cosmology

Control allows efficient exploration of complex problem spaces

Methods Used: Reinforcement learning (RL) and surrogate models

Example Applications: Accelerator control at Fermi Lab, self annealing control, light source control

Design solves optimization problems with simulations steered by machine learning (ML) and optimal experimental design methods

Methods Used: Bayesian optimization, message passing neural networks, Reinforcement learning.

Example Applications: molecular design

Inverse problems use machine learning (ML) methods to projection from observation to original form

Methods Used: Transfer Learning, Multitask Networks, Convolutional Autoencoder

Example Applications: predict material structures from X-ray or neutron scattering profiles.



ExaLearn performance portability observations and lessons learned as we move to exascale computing platforms

- There are multiple DL frameworks used across 4 application pillars: PyTorch, TensorFlow, LBANN
 - All rely on vendor libraries for accelerated linear algebra, and low-latency high-bandwidth communication
 - Vendor-specific DL hardware optimization requires different algorithmic trade-offs (packing, tensor ordering, precision)
 - Coupling and interaction between portable libraries (i.e. MPI) and accelerators is clumsy at best – vendor specific libraries are at very different levels of maturity (NCCL, RCCL)
- GPU-initiated communication is critical issue for scalable deep learning tools
- Efficient interleaving of accelerated computation streams and communication hardware is necessary to avoid stalling GPUs
- Surrogate modeling problems require data sets with large samples that drive demands on both GPU memory and parallel file system bandwidth
 - Requires the ability to partition data ingestion of a single sample across parallel ranks to avoid being I/O bound
- Complex training algorithms can include multiple instances of DL framework running concurrently with multiple scientific “environment” codes
 - Creates heterogeneous compute demands sharing GPUs and CPUs

Tokamak Fusion Device modeling can leverage all four ExaLearn pillars

• Surrogates

- ML-created models of conventional HPC particle-in-cell codes for plasma physics fusion dynamics (“GTC”)
- Synthetic “digital twins” enabling approximate real-time (RT) representation of large-scale HPC fusion simulations (“SGTC”)

• Control

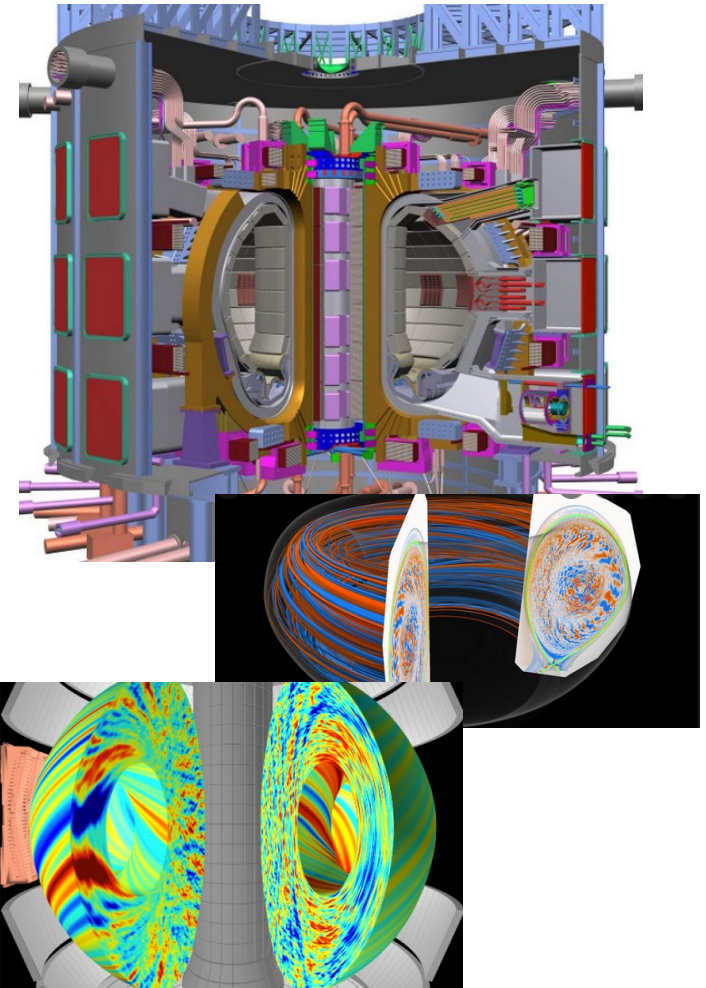
- ML-controlled fusion energy experiments with deployment in RT plasma control system (PCS) for DIII-D tokamak

• Design

- ML-generated improved future magnetic confinement fusion systems
- Optimization of proposed ideas for RT plasma behavior within “computer design space” for advanced tokamaks

• Inverse problems

- ML projection from experimental observations of plasma states to possible earlier states in evolution
- Predicting dynamical plasma structures from high-resolution diagnostic data

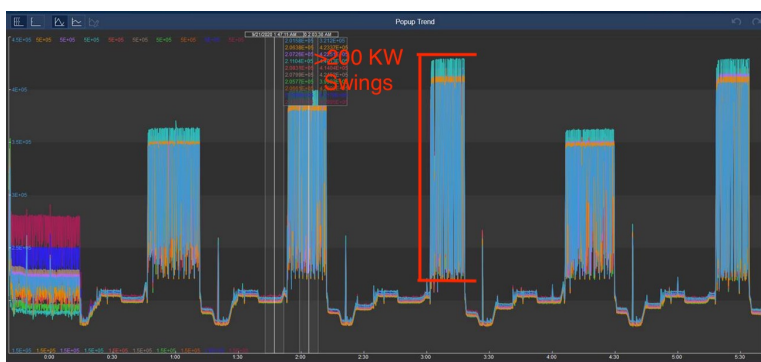


ExaLearn Investments Enabled a Gordon Bell Finalist for HPC Deep Learning Training at Scale for COVID-19 Research

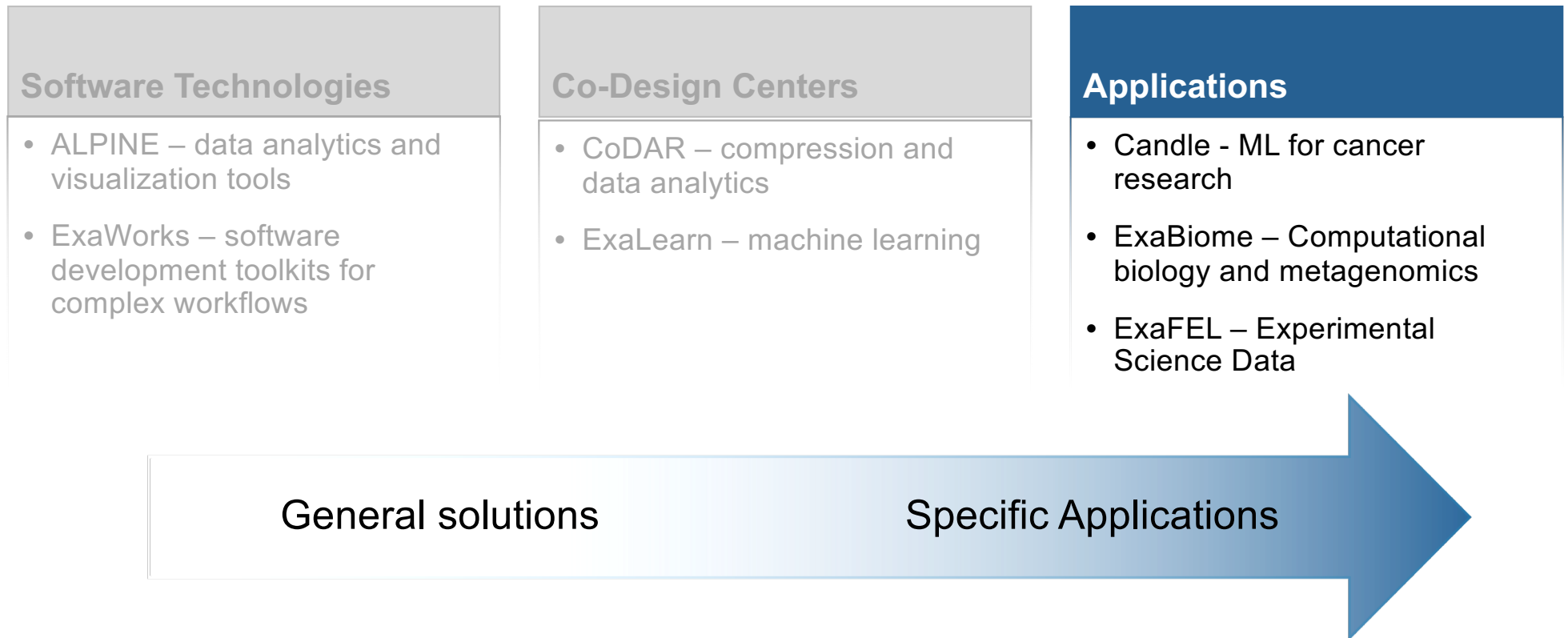
- Scaled training of molecular generator DNN for ATOM drug design loop to all of Sierra
 - Training on **1.613B molecules** achieved **318 PFLOPS** in mixed precision FP16/FP32
 - 17.1% peak half-precision efficiency
 - ~266,240 node hours over a weekend w/o LBANN SW fault
- Deep learning at scale has center-wide impact → half-precision TensorCores lead to dramatic power swings:
 - Periodic **2-3 MW swings** caused concern from power company—frequent **200 KW swings** cause center concerns
 - Asynchronous learning algorithm minimized center-wide power swings

GPUs/trainer	Trainers	Epoch time	PFLOPS
16	1040	17.2 s	253.3
8	2080	13.7 s	318.0

Table 6. Peak performance training with 4,160 nodes on Sierra.



Many ECP applications require advanced data analytics and/or machine learning to address their exascale challenge problems



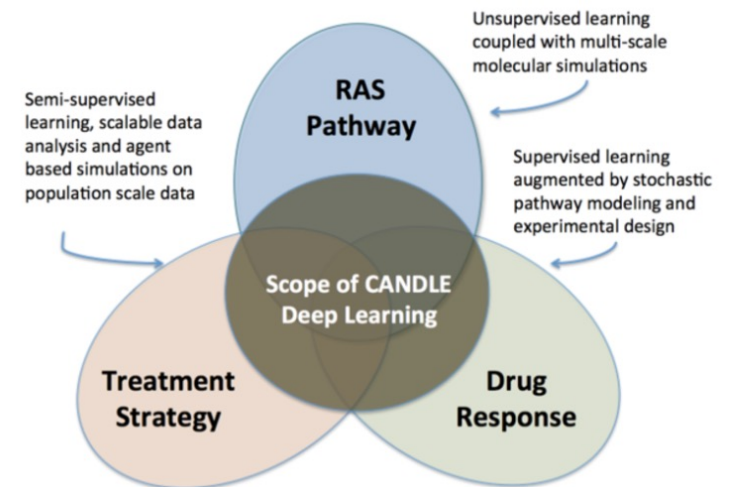
CANDLE combines the strength of exascale computing with NIH domain knowledge to address cancer and precision medicine

Science Goals:

- **Drug Response:** Predict drug response for 1000 drugs and 500,000 drug combinations on 10,000 samples using **supervised learning** to capture nonlinear relationships between drug properties and tumor properties
- **RAS Pathway:** Predict multi-scale molecular dynamics simulation state transitions for 4000 simulations and 10s of species of lipid combinations and 10s of molecular configurations using **unsupervised learning**
- **Treatment Strategy:** Predict cancer phenotypes and patient treatment trajectories using **semi-supervised learning** to read and encode millions of clinical reports

Methods:

- Exascale deep learning environment for cancer
- Use and develop open-source deep learning frameworks
- Provide open-source benchmark problems and datasets



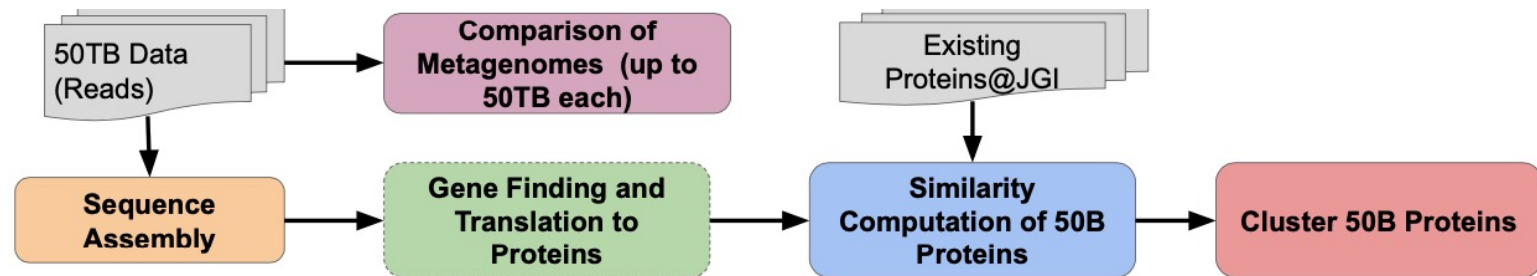
Recent results improve the confidence of using deep neural nets in these types of applications

- CANDLE framework consists of
 - DNN layer – utility library providing functions that streamline the process of writing CANDLE-compliant code; enables experiments to be designed that efficiently sweep across a range of network hyperparameters
 - Supervisor layer - provides a set of modules to enable various hyperparameter optimization (HPO) schemes and to automatically distribute the workload across available computing resources.
- Recent Results:
 - The ability to identify and understand low-confidence predictions of DNNs were significantly increased by training several thousand models on the DOE leadership computers and applying statistical methods to the outputs.
 - Combining molecular simulation and artificial intelligence on leadership-scale supercomputers—including the DOE supercomputers Summit and Theta, as well as the National Science Foundation Frontera supercomputer—is resulting in promising new insights into future COVID-19 therapeutics.
 - The CANDLE collaborative research team is applying the latest deep learning techniques for information extraction from COVID-19 and cancer-related literature. Several hundred thousand scientific reports and clinical records can be quickly and accurately scanned for relationships that shed new light on the underlying basis of diseases and provide insights toward new therapeutics.

ExaBiome develops scalable methods for metagenomic analysis

Science Goal: Compute the genome sequencing on communities with 1000s of microbial species; useful in environment, plant, animal and human health, bio-manufacturing

- Demonstrate on 50 TB of environmental data
- Challenge with methods that are not robust, evolving, and not well suited for GPUs



Application Problems

- Assemble genomes
- Compute distances using similarity networks
- Cluster computations
- Annotate and compare

Computational Motifs

- Hashtables
- Sorting
- Graph Traversal
- Generalized N-body
- Sparse matrices
- Sequence alignment

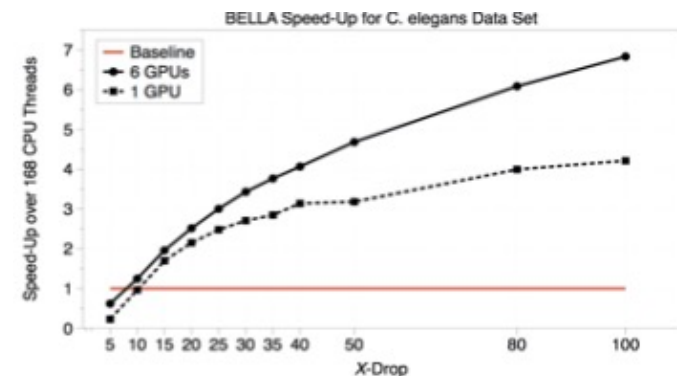
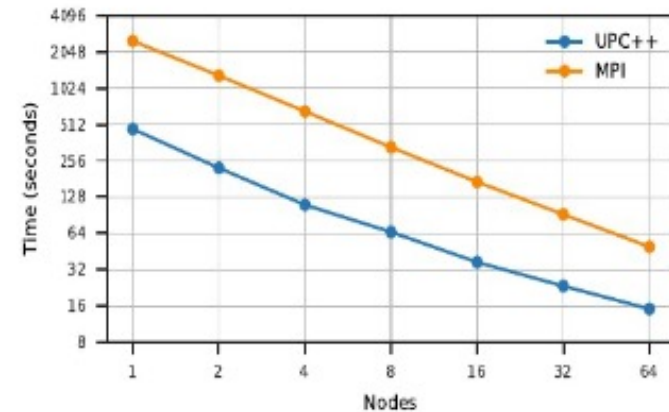


GPUs and distributed memory platforms allow for new approaches and science questions in metagenomics

- MetaHipMer Pipeline Advancements

- New parallel implementation of K-mer analysis 3-5X faster using UPC++ instead of MPI; most time consuming stage
- New scaffolding stage avoids serial bottlenecks that existed in previous versions
- Using co-assembly of all samples together (rather than one sample at a time followed by a combination step) enabled previously unachievable results

- Careful implementation of alignment kernels on GPUs enabled speed ups of 6.6X on 1 GPU and 30.7X on 6 GPUs compared to CPU implementation



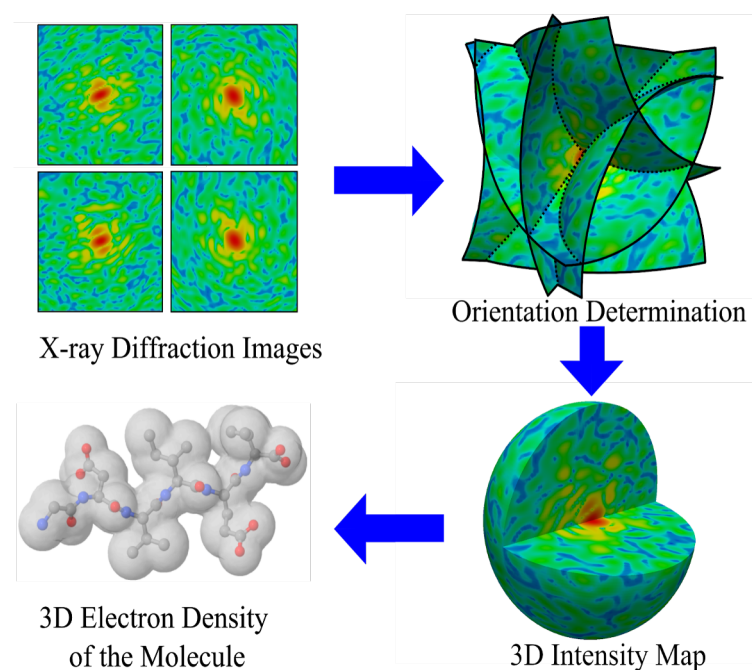
The ExaFEL project will leverage exascale computing to reduce, from weeks to minutes, the time needed to analyze LCLS data

Science Goal

- Detector data rates at light sources are advancing exponentially
- LCLS will increase data throughput by $O(10^3)$ by 2025.
- Data analysis must be carried out quickly to allow users to iterate their experiments and extract the most value from scarce beam time.

Data Challenge

- Ultrafast X-ray pulses from LCLS are used like flashes from a high-speed strobe light, producing stop-action movies of atoms and molecules
- Both data processing and scientific interpretation demand intensive computational analysis



ExaFEL: Data Analytics for High Repetition Rate Free Electron Lasers

Computational Tasks:

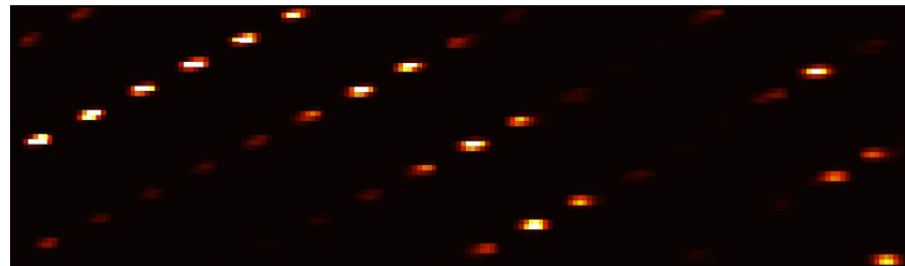
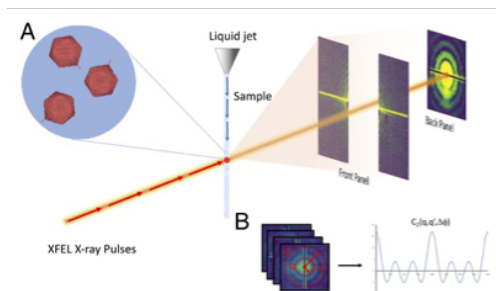
- **Serial Femtosecond Crystallography (SFX):** using x-ray tracing in nanocrystallography reconstruction (challenge problem)
- **Single Particle Imaging (SPI):** simultaneously determine conformational states, orientations, intensity, and phase from single particle diffraction images
- **Real time end-to-end workflows:** automate the coordination of resources to execute end-to-end workflows from SLAC to NERSC

Computational Challenges:

- Complex multi-component workflow, integration of DOE HPC and experimental facilities
- Moving from SFX to single particle imaging algorithms (M-TIP).
- Non-uniform FFTs on GPUs
- Improving algorithms for SFX: X-ray tracing for pixel-level resolution
- Maximum likelihood estimation optimization loop

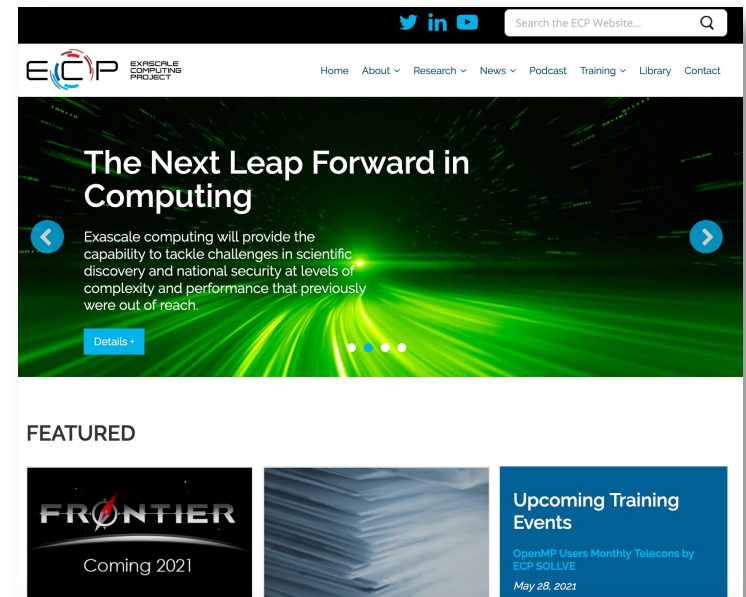
If successful, ExaFEL will enable real time analysis at LCLS and enhance the ability to answer fundamental questions about the nature of matter

- New GPU kernels for solving nanoBragg inverse problem and demonstrated on Summit using $O(10^6)$ simulated diffraction patterns. Used to ameliorate the traditional X-ray diffraction indexing algorithms for Serial Femtosecond Crystallography (SFX).
- New Cartesian/non-uniform FFT formulation of M-TIP algorithm allows SPI reconstruction to scale to many nodes; allows analysis of massive data sets expected from LCLS upgrade
- Developed the ability to stream science data from SLAC to computing facility; start analysis job on the supercomputer, and report the results back in quasi real time. Integrated with LCLS data management system.



In summary, the ECP is addressing data science challenges at the largest computing scales

- New capabilities
 - Techniques for analyzing massive scientific data sets from simulations
 - In situ methodologies
 - Better I/O, compression algorithms
 - Complex workflows
 - Scalable machine learning/deep neural nets
- Move to GPU architectures is well underway
- Technologies are used in a wide variety of applications important to the DOE
- Size of ECP project drives opportunities for increased cross team collaborations and community solutions



For more information visit
<https://www.exascaleproject.org/>

Questions?



LLNL Auspices and Disclaimer

Prepared by LLNL under Contract DE-AC52-07NA27344. This document was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor Lawrence Livermore National Security, LLC, nor any of their employees makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or Lawrence Livermore National Security, LLC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or Lawrence Livermore National Security, LLC, and shall not be used for advertising or product endorsement purposes.